# Petascale Data Storage Institute

Gary Grider, HPC-DO; Carolyn Davenport, James Nunez, HPC-5

Petascale computing infrastructures for scientific discovery make petascale demands on information storage capacity, performance, concurrency, reliability, availability, and manageability. The last decade has shown that parallel file systems can barely keep pace with high performance computing along these dimensions, posing a critical challenge when petascale requirements are considered. The LANL Petascale Data Storage Institute (PDSI) focuses on the data storage problems found in petascale scientific computing environments with special attention to community issues such as interoperability, reliability, scaling, standards, and shared tools. Leveraging the experience in applications and diverse file and storage systems expertise of its members, the institute allows a group of researchers to collaborate extensively on developing requirements, standards, algorithms, reliability, understanding, and performance tool development. Mechanisms for petascale storage and results are made available to the petascale computing community. Additionally, the institute holds periodic workshops and develops educational materials on petascale data storage for science.

**Collaborators**. The PDSI is a collaboration between researchers at Carnegie Mellon University, National Energy Research Scientific Computing Center, Pacific Northwest National Laboratory, Oak Ridge National Laboratory, Sandia National Laboratories, Los Alamos National Laboratory, the University of Michigan, and the University of California, Santa Cruz. The Los Alamos participation in the PDSI is through the LANL Institutes office via the Information Science and Technology Institute and the High Performance Computing (HPC) Division Systems Integration Group.

**Recent Accomplishments**. LANL released 9 years of computer failure data; over 23,000 records for several thousand machines. Additionally, LANL released several million usage records (job size, processors/machines used, duration, time, etc.) for the same machines for which the failure data was released. LANL also released disk failure data for a subset of these same machines. This data is publicly available and has received over 900 downloads in the last 6 months. The data has generated top paper awards at two conferences. Papers by Dr. Bianca Schroeder and Dr. Garth Gibson of Carnegie Mellon University won prizes at the 5th USENIX Conference on File and Storage Technologies (FAST 2007) and the International Symposium on Dependable Systems and Networks (DSN 2006).

The LANL data release and subsequent groundbreaking publications have encouraged an outpouring of operational data from a number of supercomputer sites and internet data and service providers. The USENIX, the Advanced

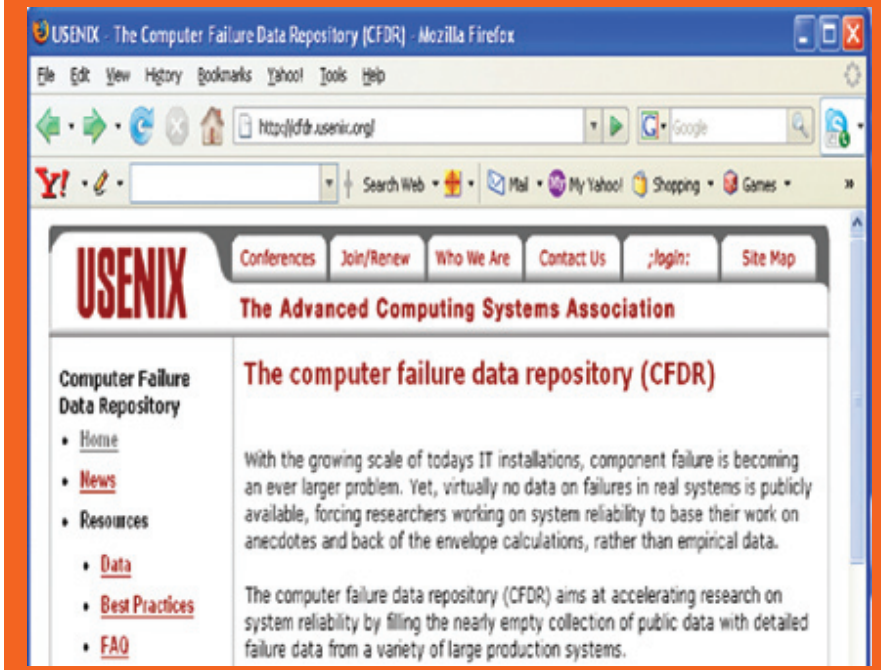


*Fig. 1. The USENIX Computer Failure Repository web site for indexing the large operational data releases to enable researchers to easily find and use these data.*

Computing Systems Association, has begun a project to index all of these large-scale operational data releases on the internet (Fig. 1). LANL is helping to organize and lead this effort.

Additionally, LANL released an important metadata synthetic benchmark, the parallel file tree walk application, which has received over 100 downloads. This is the first parallel file tree walker benchmark application of its kind. Additionally, LANL has cleared a process for releasing parallel I/O traces and file system statistics from hundreds of workstation archives and is poised to release several gigabytes of parallel I/O traces for use in the research community.

**For more information contact Gary Grider at ggrider@lanl.gov.**

**Funding Acknowledgments**
- Department of Energy, Office of Science, Scientific Discovery through Advanced Computing Program